

# **Modelling general dependence between commodity forward curves**

Mikhail Zolotko \*  
Ostap Okhrin \*

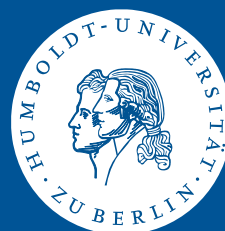


\* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>  
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin  
Spandauer Straße 1, D-10178 Berlin



# Modelling general dependence between commodity forward curves<sup>☆</sup>

Mikhail Zolotko\*, Ostap Okhrin\*\*

---

## Abstract

This study proposes a novel framework for the joint modelling of commodity forward curves. Its key contribution is twofold. First, dynamic correlation models are applied in this context as part of the modelling scheme. Second, we introduce a family of dynamic conditional correlation models based on hierarchical Archimedean copulae (HAC DCC), which are flexible, but parsimonious instruments that capture a wide range of dynamic dependencies. The conducted analysis allows us to obtain precise out-of-sample forecasts of the distribution of the returns of various commodity futures portfolios. The Value-at-Risk analysis shows that HAC DCC models outperform other introduced benchmark models on a consistent basis.

*Keywords:* commodity forward curves, multivariate GARCH, hierarchical Archimedean copula, Value-at-Risk

*JEL:* C13, C53, Q40

---

## 1. Introduction

Futures and forward contracts play a special role in the world of energy-related instruments. First, they represent one of the most widely traded type of commodity derivatives. Second, so-called forward curves, which are formed by the futures/forward prices for a

---

<sup>☆</sup>Financial support from the Deutsche Forschungsgemeinschaft through the SFB 649 “Economic Risk” is gratefully acknowledged.

<sup>\*</sup>Ladislaus von Bortkiewicz Chair of Statistics, School of Business and Economics, Humboldt-Universität zu Berlin, Spandauer Straße 1, D-10178 Berlin, Germany. E-mail: m.zolotko@gmail.com

<sup>\*\*</sup>Corresponding author, C.A.S.E. - Center for Applied Statistics and Economics, Ladislaus von Bortkiewicz Chair of Statistics, School of Business and Economics, Humboldt-Universität zu Berlin, Spandauer Straße 1, D-10178 Berlin, Germany. E-mail: ostap.okhrin@wiwi.hu-berlin.de

particular commodity for all available maturities at a certain point in time, constitute an essential input into the pricing models of more complex energy derivatives, see Pilipovic (2007). Risk management and portfolio optimisation in the situations where the value of a market position is influenced by several futures/forward prices (possibly of several commodities) is a challenging task and requires reliable tools to model the joint dynamics of all these prices.

Though in practice there is a difference between futures and forward contracts, Pilipovic (2007) notes that due to the nature of the energy commodity market, the terms forward and futures price can be used interchangeably because they represent the same value if a technical condition, namely that the delivery and the payment dates of both contracts coincide and there is no possibility of default on either side, is satisfied. Due to this indifference between forward and futures prices, the term forward curve is often used not only in the forward price context, but also to denote a set of futures prices.

Modelling of the dynamics of forward curves is in fact a goal in itself since it can provide a risk management tool for portfolios composed of futures or forward contracts, which alone can be of much interest to practitioners. The importance of modelling of joint dynamics of several forward curves is especially stressed in Ohana (2010). Moreover, the techniques and methods used for forward curves modelling can also find their application in the pricing of more complex derivatives.

All approaches to forward curves modelling can be subdivided into two broad groups. The first group consists of theoretical stochastic models of spot price dynamics that provide a framework for pricing various commodity derivatives including futures, see, e.g. Gabillon (1991), Schwartz (1997), Eydeland and Geman (1998), Cortazar and Schwartz (2003), Pilipovic (2007), and Liu and Tang (2010) to name a few. By contrast, the models of the second group start by directly formulating the process for the futures price and analyse the forward curve as a whole, see, e.g. Cortazar and Schwartz (1994), Tolmasky and Hindanov (2002), Chantziara and Skiadopoulos (2008), etc. The theoretical framework of the latter approach was first described in Reisman (1991).

Within the second group of models, in order to define the risk factors influencing the model dynamics, many authors follow the methodology of Heath et al. (1990) and apply principal components analysis (PCA). Cortazar and Schwartz (1994) is one of the first studies where PCA is used in the commodity forward curves analysis in order to define the optimal number of factors with the purpose to simulate copper futures prices and eventually price a copper-linked note. The authors were also among the first to denote the first three factors defining the dynamics of a commodity forward curve as “level”, “steepness” and “curvature”. Similar to the names of the factors identified in earlier yield curve modelling studies, such as Litterman and Scheinkman (1991), these terms were used in a somewhat abstract sense since, e.g. “level” did not mean parallel shift and “steepness” did not correspond to any commonly used steepness measure.

Tolmasky and Hindanov (2002), Koekebakker and Ollmar (2005), and Chantziara and Skiadopoulos (2008) found that in application to the forward curve dynamics modelling, PCA works reasonably well, e.g. for copper and various oil products. Two empirical factors were shown to be enough to explain over 95% of the futures return variance. At the same time, PCA yielded relatively poor results in terms of the explained variance when applied to more than one commodity curve simultaneously. In addition, the factors became less interpretable in this case, thus limiting the applicability of PCA to the modelling of the joint dynamics of several forward curves.

The two-factor model used in Ohana (2010) defines a futures price process under the physical probability measure and does not apply PCA, but employs parametrically defined factor loadings instead. Straightforward interpretability of this model is enhanced by a good explanatory power, which is comparable with that of PCA. Unlike in Tolmasky and Hindanov (2002), the factors do not explain the mechanism of the interdependence of the forward curves, this mechanism is modelled separately instead. The combination of forward curve decomposition with econometric time series analysis represents the main innovation of Ohana (2010). Other econometric models capturing interrelations between markets had already been developed previously, see, e.g. Asche et al. (2006), Bachmeier and Griffin

(2006), Dawson et al. (2006), Grasso and Manera (2007), and Hartley et al. (2008). But despite well-elaborated specifications, econometric models in these studies used only front-month futures prices which constrained the power of the applied techniques.

Chambers and Bailey (1996), Jin (2007), Onour (2009) and Ohana (2010) insinuate that the dependencies and the dynamics of the commodity markets are rather complicated and may not be well described by the methods that are based on linear specifications only and assume elliptical distributions. In this work, we develop a new copula-based econometric model trying to capture most of the dependency. The whole procedure can be sketched as follows.

First, we calibrate the two-factor model, which transforms the the whole dataset of futures returns into only four time series of shocks thus reducing the dimensionality of the modelling problem. Then, similar to Ohana (2010), the deterministic component is extracted from the shocks series by estimating the vector autoregression model, which yields heteroskedastic and temporally as well as mutually dependent residuals, which are analysed in the next step. To this end, we introduce a multivariate GARCH model based on the hierarchical Archimedean copula, which is a flexible instrument allowing for a variety of possible dependency structures for multivariate time series. The proposed model outperforms benchmark models that are driven by multivariate Gaussian innovations in terms of the quality of the out-of-sample forecasts of the portfolio returns. Moreover, the proposed model shows robustness with respect to different capital allocation scenarios. In this study we apply a combination of several approaches to the commodity spot and futures price modelling in order to deal with specific features of the data, such as common factors, autocorrelation, heteroskedasticity and non-linear dependency. These approaches are the two-factor model of the forward curve dynamics, vector autoregression and copula-based multivariate GARCH.

The paper is organised as follows. Section 2 reviews the two-factor model of the forward curve. Section 3 provides necessary theoretical background on multivariate GARCH models including those based on hierarchical Archimedean copulae. Section 4 is a simulation study

for different data-generating processes. Section 5 presents a multi-stage empirical study. Section 6 concludes.

## 2. Two-factor model of the forward curve

Let  $F(t, T)$  be the price of a forward/futures contract with maturity  $T$  (which we assume to coincide with the last trading day), observed at time  $t$ ,  $T > t$ . While  $T$  is the maturity of a futures contract in general, let  $T_t^i$  additionally denote the maturity of the  $i$ -th nearby, i.e.  $i$ -th soonest to mature, futures observed on day  $t$ .

As mentioned above, we employ the Ohana (2010) model in the first step of our analysis. This model explicitly defines factors determining the shape of the forward curve at every moment  $t$ . It also enables decomposition of any daily forward curve movement into two shocks. The first one is the long-term shock that affects all maturities equally and is caused by factors such as new information on available reserves, change of the political situation in commodity-rich countries, etc. The second one is the short-term shock that affects the market for a limited period of time, has a more significant impact on shorter than on longer maturities and is caused by factors such as temperature change, transitory supply shortage or transportation problems. For  $t \geq 0$ , the following arbitrage-free dynamics for the futures prices of a commodity is assumed:

$$\frac{F(t, T) - F(t - \Delta, T)}{F(t - \Delta, T)} = \exp \left\{ -\mathcal{N}^{-1}k(T - t + \Delta) \right\} \delta_t^S + \delta_t^L, \quad (1)$$

where  $\Delta$  is some small time interval,  $\mathcal{N} = 252$  is the considered number of trading days in a year and  $k$  is the characteristic value of a commodity that defines the extent to which a short-term shock affects longer maturities and is assumed constant for each commodity.  $(\delta^S)_{t \geq 1}$  and  $(\delta^L)_{t \geq 1}$  are short-term and long-term shocks respectively. We decompose them as:

$$\delta_t^S = \lambda_t^S + \xi_t^S, \quad \delta_t^L = \lambda_t^L + \xi_t^L, \quad (2)$$

where  $\lambda_t$  and  $\xi_t$  are deterministic and random shocks components respectively. The dynamics of  $\lambda_t$  and  $\xi_t$  is separately modelled later using VAR and multivariate GARCH

respectively. In (1) and thereafter,  $T$ ,  $t$  and  $\Delta$  are measured in trading days, and  $\Delta$  is equal to one trading day.

Both deterministic ( $\lambda_t$ ) and random ( $\xi_t$ ) components of the shocks are in general mutually dependent. Whereas the model defined by (1) is assumed to remain unchanged for both commodities throughout the whole time period considered in the study, the mechanisms defining the behaviour of the shocks components may vary over the considered time interval.

Every day  $t$  one observes a forward curve, which consists of the prices of the first to the  $M$ -th nearby futures contracts ( $M$  prices all in all). In order to decompose the futures price returns into shocks consistently, we take into account that only the price of one and the same contract can be used to calculate a return. Let  $\hat{r}_t^i$  denote the return of the  $i$ -th nearby futures contract calculated on day  $t$ .

In order to enable the calibration of the model, the effect of the short-term shock on different maturities is assumed to vanish fast enough so that the return of the farthest contract  $\hat{r}_t^M$  is not affected by it. This validity of this assumption is confirmed by the high explanatory power of the model for both commodities, see below. Since  $\hat{r}_t^M$  is then fully attributed to the long-term shock  $\delta_t^L$ ,  $\delta_t^S$  can be expressed, e.g. from (1) written for the actual return of the first nearby futures contract. The expressions for both shocks are therefore:

$$\delta_t^L = \hat{r}_t^M, \quad \delta_t^S = \exp \left\{ \mathcal{N}^{-1} k (T_t^1 - t + 1) \right\} (\hat{r}_t^1 - \hat{r}_t^M). \quad (3)$$

Plugging (3) into (1), we obtain the following formulation for the model-implied return of the futures contract with maturity  $T_t^i$ , denoted by  $r_t^i$  and written as a function of  $k$ :

$$r_t^i(k) = \hat{r}_t^M + \exp \left\{ -\mathcal{N}^{-1} k (T_t^i - T_t^1) \right\} (\hat{r}_t^1 - \hat{r}_t^M). \quad (4)$$

In order to estimate  $k$ , we minimise the root mean square error calculated from the squared differences between the model-implied ( $r_t^i$ ) and actual ( $\hat{r}_t^i$ ) futures returns using all available information, i.e. every trading day across all available maturities for each

commodity:

$$\hat{k} = \arg \min_k \sqrt{(NM)^{-1} \sum_{t=1}^N \sum_{i=1}^M \{\hat{r}_t^i - r_t^i(k)\}^2}, \quad (5)$$

where  $N$  is the number of returns observations, which requires  $N + 1$  forward curve observations, and  $M$  is the number of considered maturities.

Taking into account that discrete daily returns are an approximation of continuous returns, it follows from (1) that the logarithmised forward curve may be represented in terms of the past shocks as in (6) or as in (7) via its slope and level defined in (8) and (9) respectively, see Ohana (2010):

$$\log F(t, T) \approx \log F(0, T) + \sum_{j=1}^t \exp \{ -\mathcal{N}^{-1}k(T - j + 1) \} \delta_j^S + \sum_{j=1}^t \delta_j^L \quad (6)$$

$$= Z(T) + \exp \{ -\mathcal{N}^{-1}k(T - t) \} X_t + Y_t; \quad t \in [0; T], \quad (7)$$

where for  $t \geq 1$ ,  $X_t$  and  $Y_t$  are defined respectively as:

$$X_t = X_0 \exp(-\mathcal{N}^{-1}kt) + \sum_{j=1}^t \exp \{ -\mathcal{N}^{-1}k(t - j + 1) \} \delta_j^S, \quad (8)$$

$$Y_t = Y_0 + \sum_{j=1}^t \delta_j^L. \quad (9)$$

Here  $X_0$  and  $Y_0$  are real deterministic numbers and  $Z(T)$  is some fixed periodic function with a 12-month cycle ensuring that the forward curve has the seasonality property observed in the real data, see section 5. In (7), the term  $Y_t$  (level) regulates the vertical shift of the curve, and the term  $X_t$  (slope) defines the exact shape of the curve since the expression  $\exp \{ -\mathcal{N}^{-1}k(T - t) \}$ , which depends only on the time to maturity, is predefined for a particular commodity curve. It is easy to see that positive values of  $X_t$  will generate a curve that declines with increasing time to maturity (the situation referred to as backwardation) and, conversely, negative values of  $X_t$  generate an upward-sloping curve (which is referred to as contango).



At this point it is worthwhile to note that although the described model can generate forward curves that exhibit the seasonality property, the presence of this property is irrelevant to subsequent analysis. Since the value of the seasonality function  $Z(T)$  is constant for every futures price for some  $T$ , it cancels out when futures price returns are calculated. Thus no seasonality term shows up in (1) and in any other expressions that directly follow from it.

Ohana (2010) also provides a fundamental theoretical interpretation of the slope  $X_t$  and the level  $Y_t$  in light of the risk factors of the spot commodity price models.  $Y_t$  corresponds to the long-term commodity price and  $X_t$  is related to the convenience yield, i.e. to the relative benefit of holding of a physical commodity as compared to taking a long futures position in this commodity.

Finally, we note that the slope and the level of each forward curve can be easily estimated from the futures price data. Without loss of generality, setting  $Y_1 = 0$ , we use (9) to derive the levels for all other days.  $X_t$  is estimated by applying (7) to some pair of maturities with equal value of the seasonal function  $Z(T)$ , e.g. the 1st and the 13th, and then subtracting one expression from the other. Assuming that  $\exp\{k(T_t^{13} - 1)/\mathcal{N}\} \approx 0$ , we obtain the following approximation for the  $X_t$ :

$$X_t \approx \exp\{\mathcal{N}^{-1}k(T_t^1 - t)\} \log\left\{\frac{F(t, T_t^1)}{F(t, T_t^{13})}\right\}. \quad (10)$$

The shocks  $\delta_t^S$  and  $\delta_t^L$ , the slopes  $X_t$  and the levels  $Y_t$  represent our time series of interest in the subsequent analysis.

### 3. Copula-based multivariate GARCH

The copula-based multivariate generalised autoregressive conditional heteroskedasticity model (C-MGARCH), which is used in our study to describe the variance-covariance structure of the random component  $\xi_t$  of the shocks vector  $\delta_t = (\delta_t^{oil,L}, \delta_t^{gas,L}, \delta_t^{oil,S}, \delta_t^{gas,S})^\top$ , was proposed by Lee and Long (2009) as a generalisation of the family of multivariate generalised autoregressive conditional models (MGARCH), such as the BEKK model of

Engle and Kroner (1995), the DCC model of Engle (2002) or the VC model of Tse and Tsui (2002).

The main goal and advantage of this generalisation is the possibility to separately model linear and nonlinear dependencies between data series. MGARCH specifies the evolution of the linear dependency mechanisms, and the nonlinear dependency of error terms is controlled by a copula.

### 3.1. Copulae

The copula was introduced in the seminal paper of Sklar (1959) and defined as a continuous function  $C: [0; 1]^d \rightarrow [0; 1]$  that satisfies the equality:

$$G(x_1, \dots, x_d) = C\{F_1(x_1), \dots, F_d(x_d)\}, \quad x_1, \dots, x_d \in \mathbb{R}, \quad (11)$$

where  $G(x_1, \dots, x_d)$  is a  $d$ -dimensional multivariate distribution, and  $G_1(x_1), \dots, G_d(x_d)$  are the respective marginal distributions. For continuous multivariate distributions the copula is unique.

In practice, a class of so-called Archimedean copulae proved to be very convenient for dependency modelling purposes due to its useful properties: Archimedean copulae can generate a large variety of dependency structures, and a closed form is available for all copulae of this class. A  $d$ -dimensional Archimedean copula has the form:

$$C(u_1, \dots, u_d) = \phi^{-1}\{\phi(u_1) + \dots + \phi(u_d)\}, \quad [u_1, \dots, u_d] \in [0, 1], \quad (12)$$

where  $\phi$  is a copula generator function. Necessary and sufficient properties of  $\phi$  are provided in McNeil and Nešlehová (2009). Some of the most widely used generator functions are the Gumbel generator  $\phi_\theta(u) = (-\log u)^\theta$  for  $x \in [0, \infty)$ ,  $\theta \in [1, \infty)$ , which generates a copula with upper tail dependence:

$$C^G(u_1, \dots, u_d; \theta) = \exp\left[-\left\{(-\log u_1)^\theta + \dots + (-\log u_d)^\theta\right\}^{1/\theta}\right], \quad (13)$$

and the Clayton generator  $\phi_\theta(u) = (x^\theta - 1)/\theta$  for  $x \in [0, \infty)$ ,  $\theta \in (-1/(d-1), \infty)$ ,  $\theta \neq 0$ , which generates a copula with lower tail dependence:

$$C^{Cl}(u_1, \dots, u_d; \theta) = (u_1^\theta + \dots + u_d^\theta - 1)^{-1/\theta}. \quad (14)$$

The concept of copula also nests the case of independence of the variables, which is described by the product copula  $\Pi(u_1, \dots, u_d) = u_1 \cdot \dots \cdot u_d$ .

Since Archimedean copulae are permutation symmetric, they can be too restrictive even in the 3-dimensional case, let alone cases of higher dimensionality. A solution to this problem is a generalisation called hierarchical Archimedean copula (HAC), which is an Archimedean copula that nests other Archimedean copulae subject to certain conditions, i.e. uses them as arguments. Some of these nested copulae, which we call subcopulae, may in turn nest copulae as well, thus forming several levels of hierarchy, see Okhrin et al. (2012b). The variety of distributions that can be described by a HAC stems from the following sources: the copula's structure, the employed generator functions and the strength of the dependence as reflected by the parameter values. There exist certain limitations as to which generator functions, i.e. which copula types, can be combined, and certain conditions should be imposed on parameters to ensure that the resulting HAC is indeed a copula, see, e.g. McNeil (2008) and Hofert (2008). In this study we use only HACs employing the same kind of generator function in all subcopulae. In this case, the necessary condition for a function to be a copula is that the parameter of any outer copula, i.e. a copula nesting another copula, should not be higher than the one of any of its subordinated (inner) copulae. In other words, the strength of the relationship described by the outer copula should not exceed any of those described by its subordinated copulae. To facilitate further notation, we will say that if, e.g.  $C(x_1, x_2, x_3) = C_a(C_b(x_1, x_2), x_3)$ , where  $C_a$  and  $C_b$  are some Archimedean copulae, then  $C_b$  links  $x_1$  and  $x_2$  with each other, and  $C_a$  links  $C_b$  and  $x_3$  with each other.

The structure of a HAC will be described as  $s = (\dots(\dots i_j \dots i_k \dots) \dots (\dots) \dots)$  or  $s = (\dots(\dots \text{"}\eta_{i_j}\text{"} \dots \text{"}\eta_{i_k}\text{"} \dots) \dots (\dots) \dots)$ , where  $i_\ell \in \{1, \dots, d\}$  are the indices of the variables following the  $d$ -dimensional distribution whose dependency structure is defined by the HAC and  $\text{"}\eta_{i_\ell}\text{"}$  are the names of these variables. Every expression in brackets describes the structure of a subcopula by showing what variables and/or other subcopulae it links with each other. Figure 1 illustrates the first type of this notation demonstrating

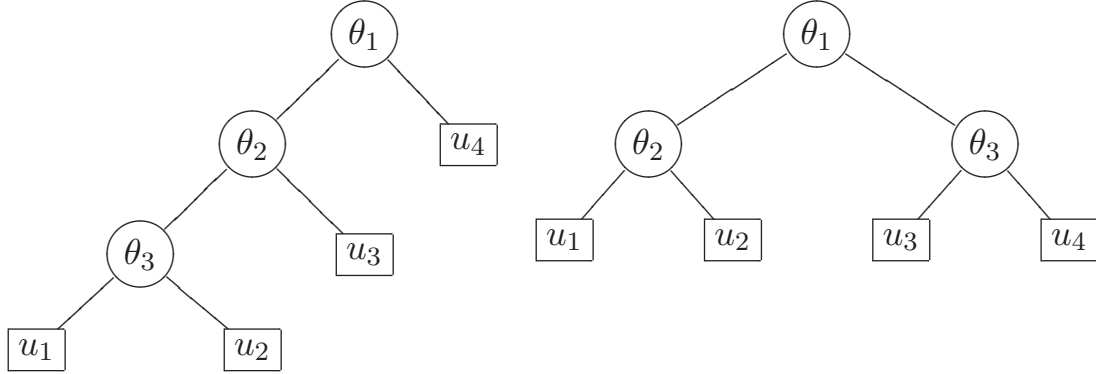


Figure 1: Graphical representation of two HAC structures for the dimension  $d = 4$ : ( $s = ((12)3)4$ ) (left) and ( $s = (12)(34)$ ) (right)

two possible HAC structures.

### 3.2. HAC-MGARCH

If  $\xi_t$ , which is a random part of the shocks vector  $\delta_t$ , is a  $d$ -variate vector with  $\mathbb{E}(\xi_t|\mathcal{F}_{t-1}) = 0$ , where  $\mathcal{F}_{t-1}$  is the information set available at  $t - 1$ , MGARCH models describe the conditional dynamics of its variance-covariance matrix  $H_t \stackrel{\text{def}}{=} \mathbb{E}(\xi_t \xi_t^\top | \mathcal{F}_{t-1})$ . It is natural to measure time in trading days as in section 2, since daily data will be analysed in the empirical part of the study.

Vector  $\xi_t$  is expressed as  $\xi_t = H_t^{1/2} e_t$  with  $e_t = \Sigma^{-1/2} \eta_t$  and  $\eta_t$  being a random vector with  $\mathbb{E}(\eta_t | \mathcal{F}_{t-1}) = 0$  and  $\mathbb{E}(\eta_t \eta_t^\top | \mathcal{F}_{t-1}) = \Sigma$ , whose components are in general nonlinearly dependent. The exact form of this dependence is described as  $\eta_t | \mathcal{F}_{t-1} \sim F_{1,\dots,d}(\eta_t, \theta) \stackrel{\text{def}}{=} C\{F_1(\eta_{1,t}), \dots, F_d(\eta_{d,t}); \theta\}$ , where  $C$  is a copula,  $\theta$  is its parameter vector assumed to be constant and  $F_1(\eta_{1,t}), \dots, F_d(\eta_{d,t})$  are marginal distributions of  $\eta_{1,t}, \dots, \eta_{d,t}$ . In the following, we assume  $F_1(\eta_{1,t}), \dots, F_d(\eta_{d,t})$  to be standard normal distributions. Since  $\theta$  is assumed to be constant, the multivariate distribution of  $\eta_t$  is not affected by any past information  $\mathcal{F}_{t-1}$ .

As mentioned earlier, several MGARCH models were proposed in the literature. In our study we focus on two of them: Dynamic conditional correlation (DCC) proposed in Engle

(2002) and the more recent Dynamic equicorrelation model (DECO), see Engle and Kelly (2012).

In DCC,  $H_t$  is modelled as  $H_t = D_t R_t D_t$ , where  $R_t = \text{diag}(Q_t^{-1/2}) Q_t \text{diag}(Q_t^{-1/2})$ ,  $Q_t = (1 - a - b)\bar{Q} + a(\varepsilon_{t-1}\varepsilon_{t-1}^\top) + bQ_{t-1}$ ,  $D_t^2 = \text{diag}(\sigma_{1,t}^2, \dots, \sigma_{d,t}^2)$ ,  $\varepsilon_t \stackrel{\text{def}}{=} D_t^{-1}\xi_t$  and  $\bar{Q}$  is the unconditional covariance matrix of  $\varepsilon_t$ ,  $a \geq 0$ ,  $b \geq 0$ ,  $a + b < 1$ . The dynamics of the individual series variances  $\sigma_{1,t}^2, \dots, \sigma_{d,t}^2$  has to be specified separately, we assume each of them to follow univariate GARCH(1,1):

$$\sigma_{\ell,t}^2 = \omega_\ell + \alpha_\ell \sigma_{\ell,t-1}^2 + \beta_\ell \zeta_{\ell,t-1}^2, \quad (15)$$

where  $\zeta_t = \sigma_t z_t$ ,  $z_t \sim \mathbf{N}(0, 1)$ ,  $\omega_\ell > 0$ ,  $\alpha_\ell \geq 0$ ,  $\beta_\ell \geq 0$  and  $\alpha_\ell + \beta_\ell < 1$ , for  $\ell = 1, \dots, d$ .

The DECO model is based on DCC though the two models are not nested. DECO applies the same decomposition of  $H_t$  as DCC, but the correlation matrix  $R_t^{DECO}$  is a transformation of  $R_t^{DCC}$ :

$$R_t^{DECO} = \begin{pmatrix} 1 & \dots & \rho_t \\ \vdots & \ddots & \vdots \\ \rho_t & \dots & 1 \end{pmatrix},$$

where  $\rho_t = 1 / \{d(d-1)\} (\iota^\top R_t^{DCC} \iota - d)$  and  $\iota$  is a unit vector with length  $d$ .

Another specification of DECO also described in Engle and Kelly (2012) is Block DECO (BDECO), which allows for a non-homogeneous structure of the vector  $\xi_t$ . Similar to the Archimedean copula, which can allow for various degrees of strength of the (nonlinear) dependence between different data series, Block DECO can accommodate different intragroup correlations within certain groups of series as well as different intergroup correlations. In the general case, if the series can be broken down into  $V$  groups of sizes  $d_1, \dots, d_V$ , the correlation matrix  $R_t$  will contain  $V$  diagonal blocks with sides equal to  $d_1, \dots, d_V$  respectively, and  $V(V-1)/2$  off-diagonal blocks:

$$R_t = \begin{pmatrix} (1 - \rho_{1,1,t})I_{d_1} & 0 & \dots \\ 0 & \ddots & 0 \\ \vdots & 0 & (1 - \rho_{V,V,t})I_{d_V} \end{pmatrix} + \begin{pmatrix} \rho_{1,1,t}J_{d_1} & \rho_{1,2,t}J_{d_1 \times d_2} & \dots \\ \rho_{2,1,t}J_{d_2 \times d_1} & \ddots & \\ \vdots & & \rho_{V,V,t}J_{d_V} \end{pmatrix}, \quad (16)$$

where  $\rho_{l,l,t} = 1/\{d_l(d_l - 1)\} \sum_{i \in l, j \in l, i \neq j} \frac{q_{i,j,t}}{\sqrt{q_{i,i,t}q_{j,j,t}}}$ ,  $\rho_{l,m,t} = 1/(d_l d_m) \sum_{i \in l, j \in m} \frac{q_{i,j,t}}{\sqrt{q_{i,i,t}q_{j,j,t}}}$ ,  $\rho_{l,m,t} = \rho_{m,l,t}$  for all  $l, m$  and  $q_{i,j,t}$  is the  $i, j$ -th element of the matrix  $R_t^{DCC}$ , so BDECO correlations are in fact average correlations of each block of the matrix  $R_t$  in the DCC model.

To denote the structure of BDECO we use the same notation that was introduced for the copula structure in section 3.1. Each expression in brackets refers in this case to one of the  $V$  groups of variables.

### 3.3. Estimation of copula-based multivariate GARCH

The assumptions on the distribution of  $\eta_t$  give rise to the following log-likelihood function which should be maximised for model parameter estimation (see Lee and Long (2009)):

$$\mathcal{L}(\varphi, \theta) = \sum_{t=1}^N \log f_{1,\dots,d}(\eta_t; \theta) + \log |\Sigma^{1/2}(\theta) H_t^{-1/2}(\varphi)|, \quad (17)$$

where  $N$  is the number of observations of vector  $\xi_t$ ,  $\varphi$  is the vector of MGARCH parameters defined as  $\varphi = (\omega_1, \dots, \omega_d, \alpha_1, \dots, \alpha_d, \beta_1, \dots, \beta_d, a, b)^\top$  for both DCC and DECO models and  $f_{1,\dots,d}(\eta_t; \theta)$  is the pdf of  $\eta_t$ . The log-likelihood function is to be maximised over  $\varphi$  and  $\theta$ :

$$(\hat{\varphi}, \hat{\theta}) = \arg \max_{\varphi, \theta} \mathcal{L}(\varphi, \theta). \quad (18)$$

With the HAC assumption on the distribution of  $\eta_t$  taken into account, the log-likelihood function (17) for HAC-MGARCH takes the form:

$$\begin{aligned} \mathcal{L}(\varphi, \theta) = \sum_{t=1}^N \left[ \sum_{i=1}^d \{\log f_i(\eta_{i,t})\} + \log c\{F_1(\eta_{1,t}), \dots, F_d(\eta_{d,t}); \theta\} + \right. \\ \left. + \log |\Sigma^{1/2}(\theta) H_t^{-1/2}(\varphi)| \right], \end{aligned} \quad (19)$$

where  $f_i(\eta_{i,t})$  is the marginal pdf of  $\eta_{i,t}$  assumed to be the pdf of standard normal distribution and  $c(u_1, \dots, u_d) = \frac{\partial^d C(u_1, \dots, u_d)}{\partial u_1 \dots \partial u_d}$  is the copula density function.

Within one MGARCH type, e.g. DCC or VC, HAC-MGARCH represents the most general model. AC-MGARCH is thus its special case, which in turn nests standard MGARCH.

Since standard MGARCH assumes independence of  $\eta_{i,t}$ ,  $C(\cdot)$  takes the form of the product copula  $\Pi(\cdot)$  in this case, hence  $\log c(\cdot)$  turns to zero and  $\Sigma^{1/2}$  becomes a unity matrix.

In the more general case of AC-MGARCH and HAC-MGARCH, the diagonal elements of  $\Sigma$  are equal to 1 since  $\eta_{i,t}$  follows standard normal distribution with unity variance. The off-diagonal elements of  $\Sigma$  can be evaluated using Hoeffding's lemma (Hoeffding (1940)). According to it, if  $\eta_1$  and  $\eta_2$  are random variables with the marginal distributions  $F_1$  and  $F_2$  and the joint distribution  $F_{12}$ , and if their first and second moments are finite, then taking Sklar's theorem into account:

$$\begin{aligned}\sigma_{12}(\theta) &= \iint_{\mathbb{R}^2} \{F_{12}(\eta_1, \eta_2; \theta) - F_1(\eta_1) F_2(\eta_2)\} d\eta_1 d\eta_2 = \\ &= \iint_{\mathbb{R}^2} [C\{F_1(\eta_1), F_2(\eta_2); \theta\} - F_1(\eta_1) F_2(\eta_2)] d\eta_1 d\eta_2.\end{aligned}\tag{20}$$

#### 4. Simulation study

This brief simulation study demonstrates the performance of HAC-MGARCH in the cases of correct and incorrect model specifications. To be consistent with the empirical study, we consider a 4-dimensional case.

We simulate data series with length  $N = 2000$  from standard DCC, Gumbel HAC DCC with  $s = (12)(34)$ , as in Figure 1 (right panel), and BDECO with  $s = (12)(34)$ .<sup>1</sup> Each of the simulated samples is then used to estimate the following models: DCC, Gumbel AC DCC, Gumbel HAC DCC, DECO and BDECO. We include HAC DCC and BDECO to be able to consider a typical situation in practice where portfolio components can be naturally grouped in a certain way, anticipating the relevance of this situation for our case.

The vectors of the univariate GARCH parameters for the four series are  $\omega = (0.013, 0.003, 0.014, 0.016)^\top$ ,  $\alpha = (0.15, 0.2, 0.06, 0.2)^\top$ ,  $\beta = (0.6, 0.7, 0.8, 0.6)^\top$ . The DCC parameters are  $a = 0.1$  and  $b = 0.7$ . The chosen values are in line with those encountered in

---

<sup>1</sup>We also considered DECO and Gumbel AC DCC. The corresponding results are not reported here for the sake of brevity, but are available upon request.

practice:  $\alpha^j$  and  $a$  estimates are usually relatively close to their lower boundary 0, whereas  $\beta^j$  and  $b$  estimates are often close to their higher boundary 1. The copula parameters are  $\theta \stackrel{\text{def}}{=} (\theta_1, \theta_2, \theta_3)^\top = (1.5, 3.0, 5.0)^\top$  for HAC DCC, where  $\theta_1$  is the parameter of the outer copula and  $\theta_2$  and  $\theta_3$  are the parameters of the two subordinated copulae, see Figure 1 (right panel).

The whole procedure is repeated  $J = 1000$  times. To quantify the quality of fit, we use the conditional Kullback-Leibler divergence (Kullback and Leibler (1951)):

$$KL(\xi_t, \psi_t, \psi_t^*) = \mathbb{E}_{\psi_t} \left\{ \log \frac{\psi_t(\xi_t, \varphi, \theta)}{\psi_t^*(\xi_t, \hat{\varphi}^*, \hat{\theta}^*)} \right\}, \quad (21)$$

where  $\psi_t(\xi_t, \varphi, \theta)$  is the true conditional density of the observation  $\xi_t$  calculated using the true parameter vector values  $(\varphi, \theta)^\top$  and  $\psi_t^*(\xi_t, \hat{\varphi}^*, \hat{\theta}^*)$  is the conditional density of the observation  $\xi_t$  according to the candidate model calculated using the estimated parameter vector  $(\hat{\varphi}^*, \hat{\theta}^*)^\top$ . We estimate  $KL$  by taking the average of the logarithms of the actual calculated conditional probability ratios:

$$\widehat{KL}(\xi_t, \psi_t, \psi_t^*) = \frac{1}{N} \sum_{t=1}^N \left\{ \log \frac{\psi_t(\xi_t, \varphi, \theta)}{\psi_t^*(\xi_t, \hat{\varphi}^*, \hat{\theta}^*)} \right\}. \quad (22)$$

The kernel density of  $\widehat{KL}$ , shown in Figures 2-4, is calculated using the Gaussian kernel smoother. The bandwidth is chosen based on Silverman's rule of thumb.

Smaller absolute values of  $\widehat{KL}$  correspond to a model that is closer to the true one. Estimation of a true specification results in  $\widehat{KL}$  distributed closely around zero, corresponding to the highest possible likelihood.

Figure 2 shows the distribution of  $\widehat{KL}$  based on 5 specifications, where the true model is standard DCC. Not surprisingly, the lines corresponding to standard DCC, AC DCC and HAC DCC are indistinguishable from each other since standard DCC is a special case of the other two models. The lines corresponding to both DECO specifications are positioned relatively far from those corresponding to the DCC specifications, suggesting that the DECO models' ability to approximate DCC models is rather low.

The results of estimation of the dataset generated by HAC DCC are shown in Figure 3. It is obvious that the simpler DCC models fail to capture the dependence structure gener-



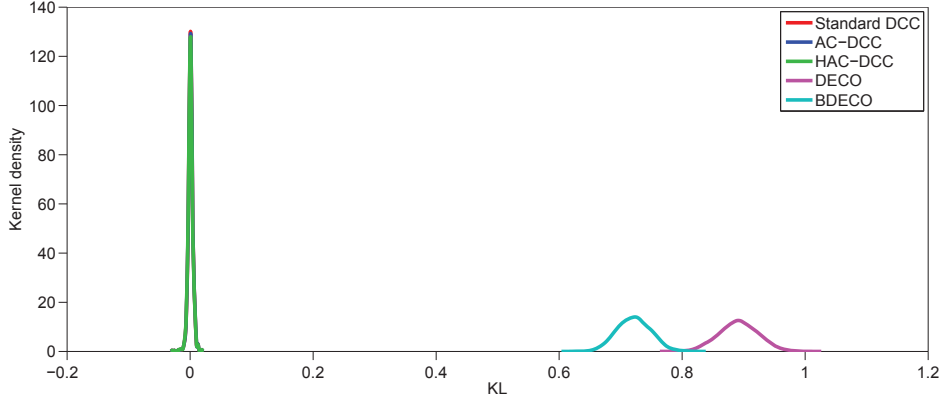


Figure 2:  $\widehat{KL}$  kernel density estimation under various MGARCH specifications, true model: standard DCC. Assumed specifications: standard DCC, Gumbel AC DCC, Gumbel HAC DCC with  $s = (12)(34)$  (last three lines coincide), DECO, BDECO with  $s = (12)(34)$ .

ated by the more general models. At the same time, the dependence assumed by HAC DCC can be satisfactorily approximated only by HAC DCC itself. DECO models are generally a rather poor fit for HAC DCC.

BDECO is the model that generated the dataset analysed in Figure 4. One sees that the  $\widehat{KL}$  distributions resulting from the true BDECO and all DCC models estimation are reasonably close to each other, which means that all DCC models perform equally in this case and all significantly outperform DECO. It is noteworthy that the distributions corresponding to the DCC models in Figure 4 are positioned much closer to zero than the distributions corresponding to DECO and BDECO in Figures 2 and 3, which can be interpreted as a better ability of DCC to approximate DECO than the ability of the latter to approximate the former.

To summarise, we can conclude that HAC DCC approximates all other models under consideration fairly well, whereas the opposite is not true. This implies that the HAC DCC specification in fact looks promising since it can capture dependence structures that other models cannot, at the same time it can approximate dependence structures defined by other models well enough.

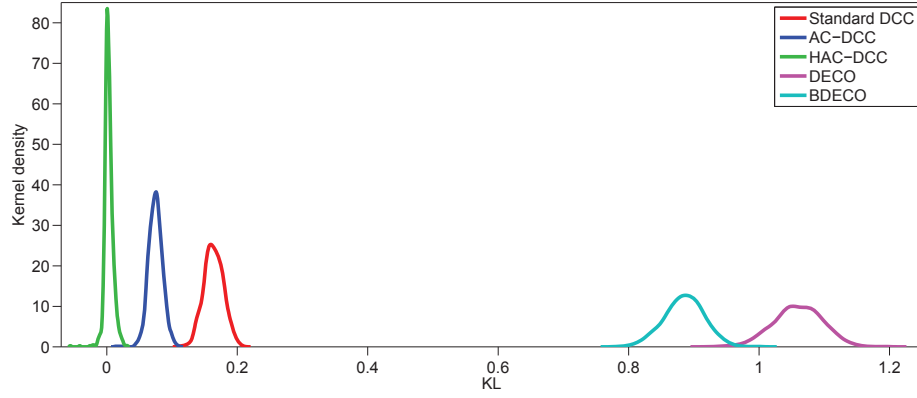


Figure 3:  $\widehat{KL}$  kernel density estimation under various assumed MGARCH specifications, true model: DCC with Gumbel HAC, ( $s = (12)(34)$ ). Assumed specifications: standard DCC, Gumbel AC DCC, Gumbel HAC DCC with  $s = (12)(34)$ , DECO, BDECO with  $s = (12)(34)$ .

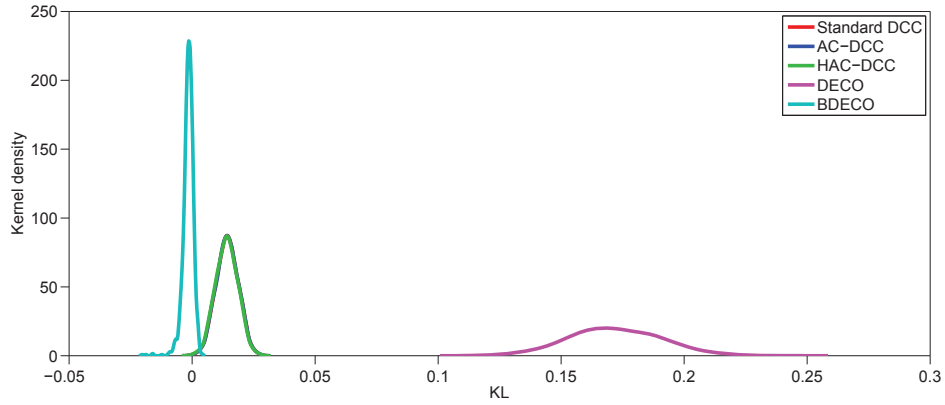


Figure 4:  $\widehat{KL}$  kernel density estimation under various assumed MGARCH specifications, true model: BDECO, ( $s = (12)(34)$ ). Assumed specifications: standard DCC, Gumbel AC DCC, Gumbel HAC DCC with  $s = (12)(34)$  (last three lines almost coincide), DECO, BDECO with  $s = (12)(34)$ .

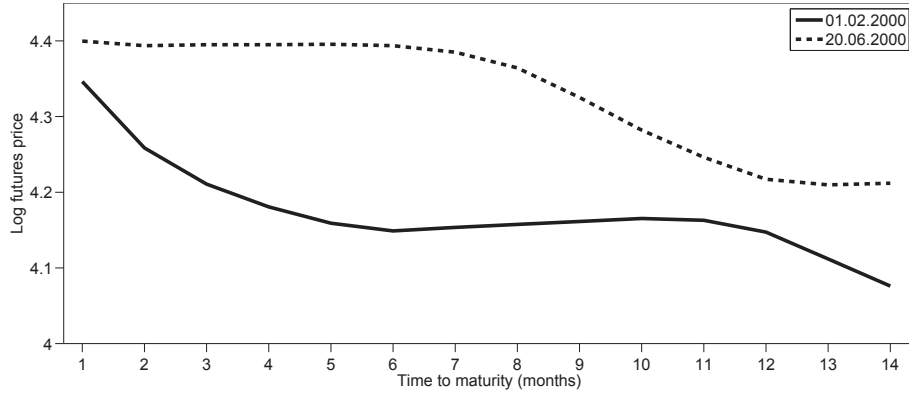


Figure 5: Logarithmised heating oil forward curves on 01.02.2000 and 22.06.2000 in US cents per gallon

## 5. Empirical study

In this study we use daily New York Harbour No. 2 Heating Oil Futures and Henry Hub Natural Gas Futures prices (obtained from Bloomberg) for the time frame from 01.02.2000 to 19.12.2011 for the first  $M = 14$  maturities, which resulted in  $N = 2975$  observations for each of the two forward curves. Figure 5 demonstrates a logarithmised heating oil forward curve for 01.02.2000 and 22.06.2000 and exposes its two important features. The first one is backwardation, i.e. a situation where more distant maturities are traded at lower prices than less distant ones (both curves are decreasing). This is a normal situation in commodity markets as stated, e.g. in Gabillon (1991) and reflects (buyers’) “preference for the present time whatever the reasons are”. The other feature is seasonality: contracts expiring in winter are generally more expensive than those expiring in summer. As discussed in section 2, the seasonality effect is of no consequence to our analysis.

### 5.1. Estimation of the Model

In the first step, we calibrate the two-factor model (1). The calibrated parameters are  $\hat{k}_{oil} = 3.10$  and  $\hat{k}_{gas} = 2.86$ , and the model explains 98.23% and 93.80% of the variance of the heating oil and natural gas futures returns, respectively, which can be regarded as a good fit. Due to a longer time span, our results are slightly different from those obtained by

Series	ADF	KPSS	Series	ADF	KPSS
$Y^{oil}$	-1.43 (0.55)	4.82 (0.01)	$\delta^{oil,L}$	-19.80 (0.00)	0.17 (0.10)
$Y^{gas}$	-1.66 (0.45)	1.96 (0.01)	$\delta^{gas,L}$	-19.03 (0.00)	1.05 (0.01)
$X^{oil}$	-3.07 (0.03)	2.84 (0.01)	$\delta^{oil,S}$	-19.91 (0.00)	0.15 (0.10)
$X^{gas}$	-3.71 (0.00)	1.85 (0.01)	$\delta^{gas,S}$	-20.34 (0.00)	0.05 (0.10)

Table 1: Test statistics of the stationarity tests, p-values in brackets.

Ohana (2010), i.e. we find evidence for an even higher explanatory power of the model for heating oil and a little lower explanatory power for natural gas. Given  $\hat{k}$ , we estimate shocks series  $\delta_t$ , levels  $Y_t$  and slopes  $X_t$  for both commodities, see Figures 6, 7 and 8 respectively.

Results of some basic time series analysis are provided in Table 1. As can be concluded from it, the Augmented Dickey-Fuller (ADF) test does not reject the unit root hypothesis for the levels series, and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test does not reject the stationarity hypothesis for the shocks series. Although the tests disagree about the stationarity of the slopes series, we decided to use them as regressors in the model in line with Ohana (2010).

## 5.2. Vector autoregression

After obtaining the shocks series  $\delta_t = \left( \delta_t^{oil,L}, \delta_t^{gas,L}, \delta_t^{oil,S}, \delta_t^{gas,S} \right)^\top$ , Ohana (2010) recommends estimating a vector error correction model (VECM) as the levels of the two commodities  $Y_t^{oil}$  and  $Y_t^{gas}$  are cointegrated. Following the approach of Engle and Granger (1987), the long-term relationship between the commodity levels is estimated in a separate model and turns out to be well approximated by a piecewise linear function, see Figure 9. In the next step, the residuals of this model, i.e. deviations from the long-term relationship, are used as an exogenous variable in the VAR model. In this paper we do not apply the VECM framework for two reasons. First, even with cointegration accounted for, the model in Ohana (2010) explains a too small proportion of the shocks' variance: the four equations within the VECM model explaining the dynamics of the heating oil long-term shocks  $\delta_t^{oil,L}$ , natural gas long-term shocks  $\delta_t^{gas,L}$ , heating oil short-term shocks  $\delta_t^{oil,S}$  and natural gas

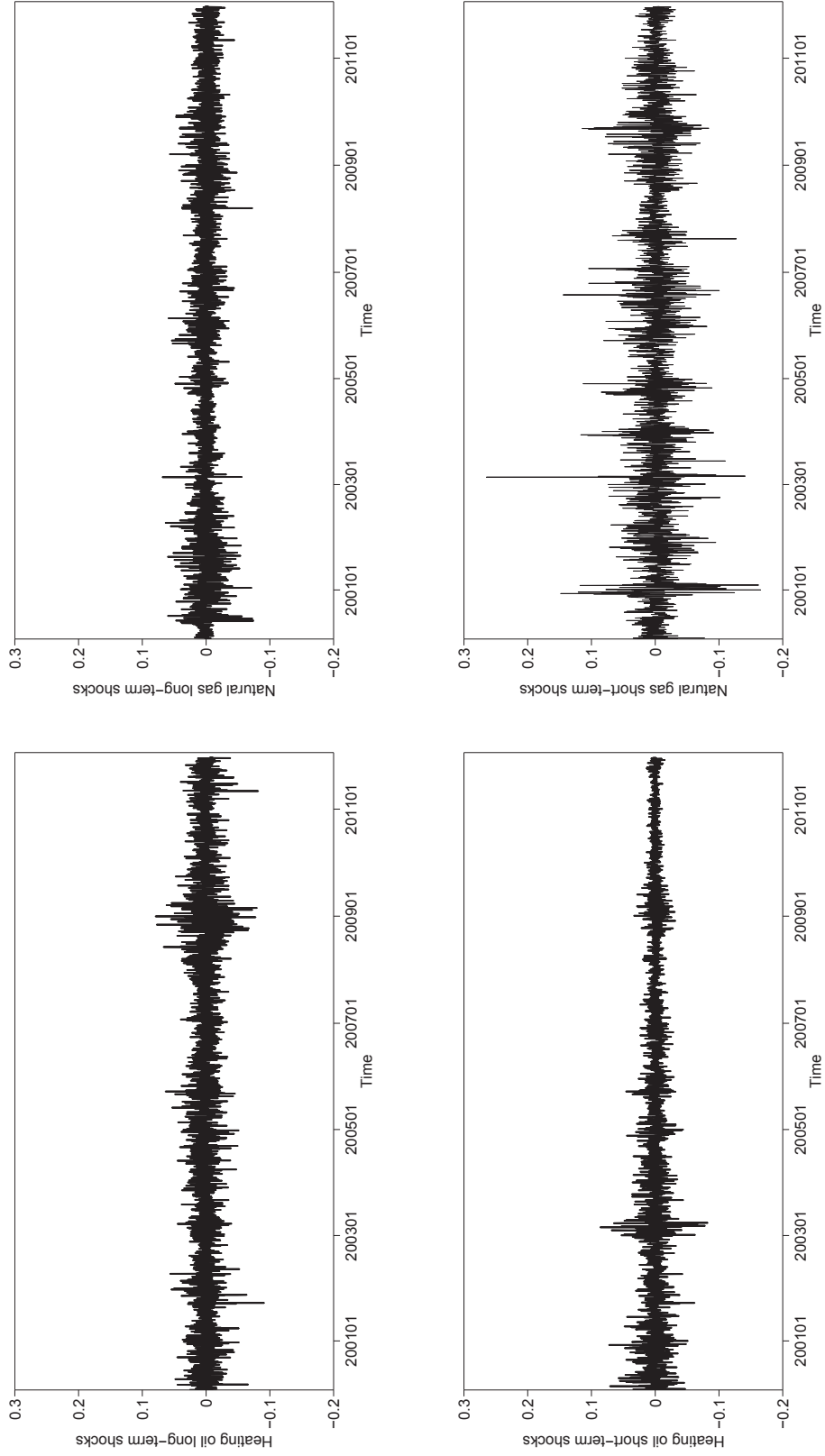


Figure 6: Estimated shocks in the two-factor model of the forward curve: heating oil long-term shocks ( $\delta^{oil,L}$ ) (upper left), natural gas long-term shocks ( $\delta^{gas,L}$ ) (upper right), heating oil short-term shocks ( $\delta^{oil,S}$ ) (lower left), natural gas short-term shocks ( $\delta^{gas,S}$ ) (lower right) (01.02.2000 – 19.12.2011).

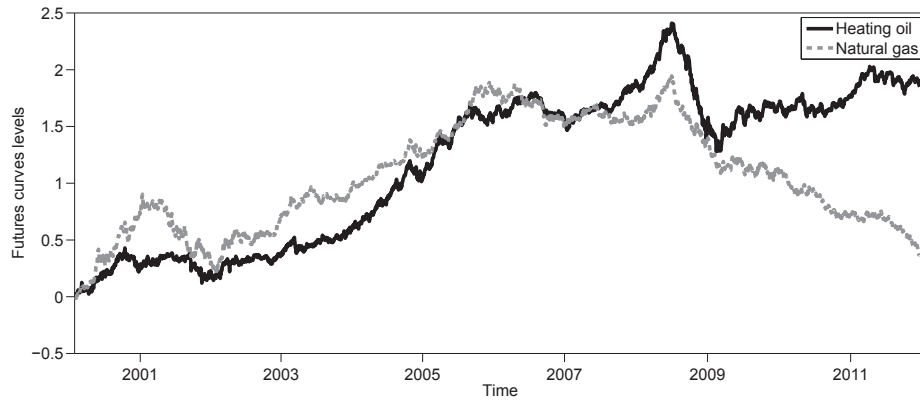


Figure 7: Heating oil ( $Y_t^{oil}$ ) and natural gas ( $Y_t^{gas}$ ) levels (01.02.2000 – 19.12.2011).

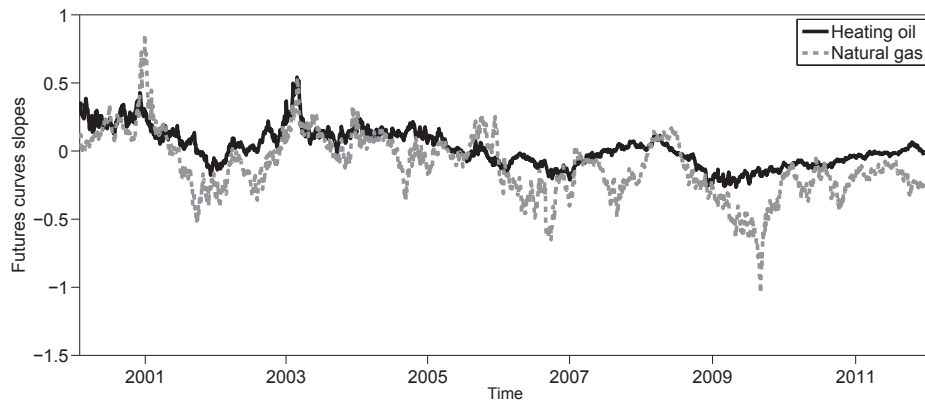


Figure 8: Heating oil ( $X_t^{oil}$ ) and natural gas ( $X_t^{gas}$ ) slopes (01.02.2000 – 19.12.2011).

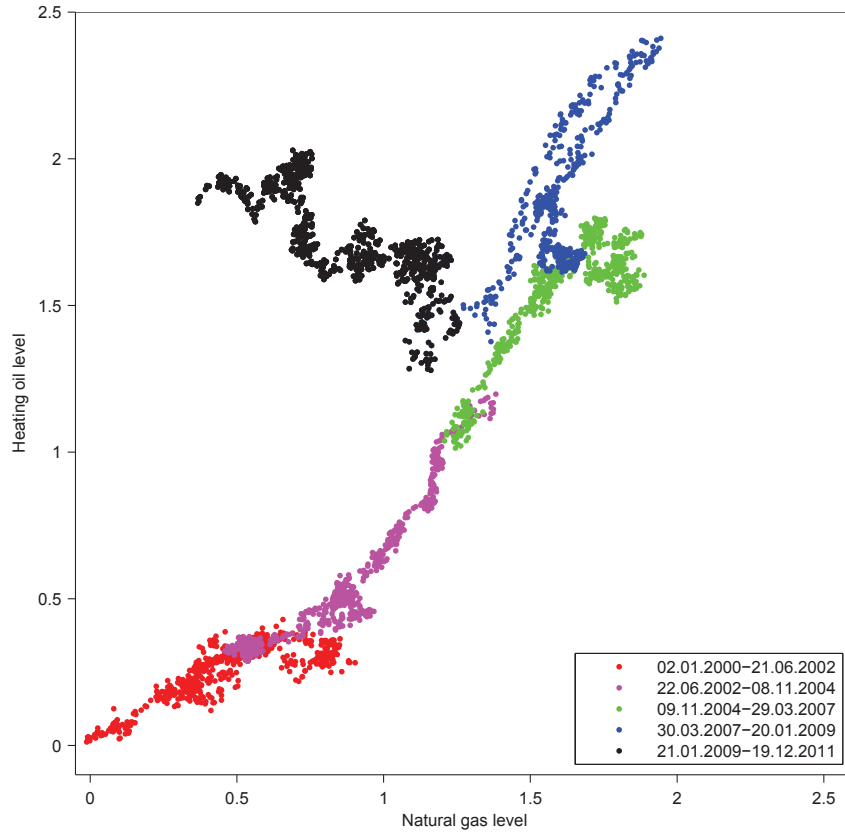


Figure 9: Long-term relationship between heating oil ( $Y_t^{oil}$ ) and natural gas ( $Y_t^{gas}$ ) levels

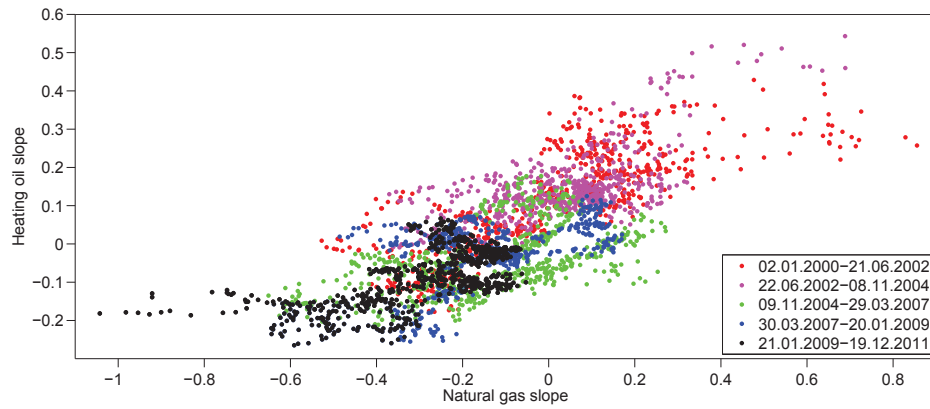


Figure 10: Long-term relationship between heating oil ( $X_t^{oil}$ ) and natural gas ( $X_t^{gas}$ ) slopes

short-term shocks  $\delta_t^{gas,S}$  had  $R^2$  of only 3.19%, 1.89%, 2.21% and 1.72%, respectively. Second, there is enough evidence of the fact that during 2009 the cointegration link between oil and gas was substantially affected by the changes in the US gas market. A quick look at Figure 9 confirms this: whereas oil prices have recovered after their slump during the 2008-2009 crisis, gas prices have not. De Bock and Gijón (2011) cite an additional supply of non-conventional gas (above all shale gas) and high storage levels in the US market as the main reasons behind the relative weakness of US gas prices and the loosening of the link between the prices for natural gas and West Texas Intermediate (WTI), which is a low-sulfur light grade of oil and one of the world's benchmarks in oil pricing shown to be cointegrated with heating oil, see, e.g. Hartley et al. (2008). This period can also be denoted as a period of increased uncertainty in the gas market because it appears to be difficult to reliably estimate the amount of the gas that can be extracted. Nevertheless, it is reasonable to assume that the long-term link between heating oil and natural gas prices will not be completely broken because there seems to be little reason for the common factors on the demand side and for the link between the two commodities to disappear. This is why the period that started in 2009 should be denoted as a transition to a new long-term relationship between the two commodities. As an extra argument, one can refer to the Annual Energy Outlook (2012) by the US Energy Information Administration (EIA), which makes projections for a number of energy market indicators including commodity prices and uses the virtual "low-sulfur light crude oil price", denoted as "similar" to the price for WTI, as the reference oil price. An important for us projection is the one for the ratio of low-sulfur light crude oil price to Henry Hub natural gas price, which is predicted to change very slowly after achieving what appears to be its peak around 2015 following an especially steep rise during 2007-2010. Figure 9 shows that during the period covered in Ohana (2010) (up to February 2009), the cointegration approach seemed reasonable, but the newer data, shown in black, gives some impression of a negative correlation between heating oil and natural gas price levels. We cannot treat this negative relationship as part of the long-term price interdependence because no theoretical arguments would justify this.



As a result, the model specification used in this paper is a simpler vector autoregression (VAR) with the maximum lag of 1 and two extra regressors  $X_{t-1}^{oil}$  and  $X_{t-1}^{gas}$ :

$$\delta_t = \lambda_t + \xi_t = \mu + \Gamma\delta_{t-1} + \Psi A_{t-1} + \xi_t, \quad (23)$$

where, as noted earlier,  $\delta_t = (\delta_t^{oil,L}, \delta_t^{gas,L}, \delta_t^{oil,S}, \delta_t^{gas,S})^\top$  is the shocks vector. Its deterministic component is defined as  $\lambda_t \stackrel{\text{def}}{=} \mu + \Gamma\delta_{t-1} + \Psi A_{t-1}$ , where  $\mu$  is a vector of constants,  $A_t = (X_t^{oil}, X_t^{gas})^\top$  is the slopes vector,  $\xi_t$  with  $E(\xi_t | \mathcal{F}_{t-1}) = 0$  is the random shocks component vector that follows C-MGARCH and  $\Gamma$  and  $\Psi$  are parameter matrices.

The maximum lag of 1 was chosen based on the Hannan-Quinn and Bayesian Schwartz information criteria. For both criteria, the generalisation for multivariate processes was calculated, see Lütkepohl and Krätzig (2004) for details.

The model is estimated in a rolling window fashion with a window size of 500 and a step of 5 days. We treat VAR estimation as a procedure whose only aim is to extract the deterministic component  $\lambda_t$  from the vector  $\delta_t$ . The exact form of the autoregressive relation is not as important to us as the obtained residual series  $\hat{\xi}_t$ , to which MGARCH models are applied in order to describe their dependency dynamics.

For MGARCH we consider eleven different models, among them two AC DCC specifications (with Gumbel and Clayton generators) and four HAC DCC specifications (with Gumbel and Clayton generators as well as two different structures  $s_1$  and  $s_2$  defined below). Three DECO specifications are estimated for reference purposes: DECO and BDECO with  $s_1$  and  $s_2$ . We also estimate univariate GARCH, which is effectively a special case of MGARCH and can be characterised as DCC or DECO with constant zero correlation between the components of  $\xi_t$ .

Usually, estimation of HAC-MGARCH or BDECO models implies not only estimation of the parameters, but also the choice of an optimal structure. Here we consider the following structures:  $s_1 = ((ol\ gl)(os\ gs))$  and  $s_2 = ((ol\ os)(gl\ gs))$ , where  $ol$ ,  $gl$ ,  $os$ ,  $gs$  are the components of the vector  $\eta_t$  in the MGARCH model (see section 3.3) that correspond to heating oil long-term shocks ( $\delta_t^{oil,L}$ ), natural gas long-term shocks ( $\delta_t^{gas,L}$ ), heating oil short-term shocks ( $\delta_t^{oil,S}$ ) and natural gas short-term shocks ( $\delta_t^{gas,S}$ ), respectively. The grouping

in  $s_2$  is the result of our preliminary analysis of the residuals based on the structure selection procedure from Okhrin et al. (2012a), and  $s_1$  is its natural analogue. The interpretation of both structures is straightforward.

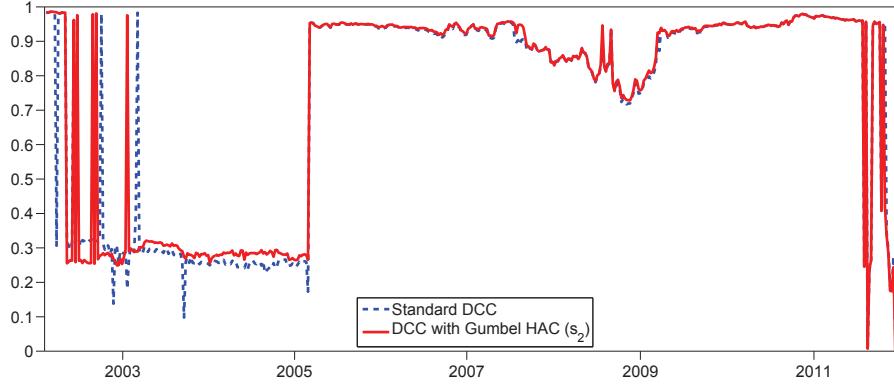


Figure 11: Parameter  $b$  in the DCC part (01.02.2000 – 19.12.2011). Values were estimated on a 500-day period ending on the respective day.

Figure 11 demonstrates the variation of the parameter  $b$  in the main DCC equation from period to period. This can be seen as a justification for using several time windows for the model estimation.

### 5.3. Value-at-Risk of the portfolio

Portfolio Value-at-Risk backtesting became a standard tool for model quality assessment. The aim of this procedure is to compare the precision of Value-at-Risk forecasts produced by different models for different futures portfolios. According to the two-factor model of the forward curve (1) the return of any portfolio consisting of heating oil and natural gas futures can be expressed as a “portfolio” of the four shocks. The heating oil shocks weights in such a portfolio  $w_{\delta_t^{oil,S}}$  and  $w_{\delta_t^{oil,L}}$  are expressed as:

$$w_{\delta_t^{oil,S}} = \sum_{i=1}^M w_i^{oil} \exp\{-\mathcal{N}^{-1}k_{oil}(T_t^i - t + 1)\}, \quad w_{\delta_t^{oil,L}} = \sum_{i=1}^M w_i^{oil}, \quad (24)$$

where  $w_i^{oil}$  are weights of  $M = 14$  heating oil contracts in the futures portfolio and other notation is as defined above. The weights of the natural gas shocks are calculated analo-

gously. Thus, if the joint distribution of the shocks can be estimated for a particular day, the distribution for any futures portfolio return, i.e. for any weighted sum of the shocks, can be easily obtained.

Let the portfolio return  $\mathcal{R}_{t+1}$  for an arbitrary day  $t + 1$  be calculated as:

$$\mathcal{R}_{t+1} = w^\top \hat{r}_{t+1} = w_s^\top \delta_{t+1}, \quad (25)$$

where  $w = (w_1, w_2, \dots, w_{28})^\top$  is the vector of the futures weights (14 maturities for each commodity),  $w_s = (w_{s1}, w_{s2}, w_{s3}, w_{s4})^\top$  is the corresponding vector of the shocks weights (the notation is changed for the sake of simplicity),  $\hat{r}_{t+1}$  is the vector of futures returns at  $t + 1$  and  $\delta_{t+1}$  is defined as in (23). Then the Value-at-Risk (VaR) at level  $0 < \alpha < 1$  for day  $t + 1$  is defined as  $VaR_{t+1}(\alpha) \stackrel{\text{def}}{=} F_{\mathcal{R}_{t+1}}^{-1}(\alpha)$ . From (2) it follows that:

$$VaR_{t+1}(\alpha) = w_s^\top \lambda_{t+1} + F_{w_s^\top \xi_{t+1}}^{-1}(\alpha). \quad (26)$$

Using the estimated VAR parameters, we calculate the forecast for  $\lambda_{t+1}$ . For non-copula-based MGARCH models, since  $e_{t+1} \sim \mathbf{N}(0, I_d)$ , the expression  $w_s^\top \xi_{t+1}$  is also normally distributed, i.e.  $w_s^\top \xi_{t+1} \sim \mathbf{N}(0, w_s^\top H_{t+1} w_s)$ . Hence  $F_{w_s^\top \xi_{t+1}}^{-1}(\alpha)$  can be inferred explicitly using properties of the normal distribution. For AC-MGARCH and HAC-MGARCH such an estimation is not possible since  $e_{t+1} = \Sigma^{-1/2} \eta_{t+1}$  is not normal, hence the whole distribution of  $H_{t+1}^{1/2} e_{t+1}$  has to be simulated based on the estimated copula parameters governing the distribution of  $\eta_{t+1}$ . For each distribution we simulate 3000 points, then obtain the distribution of the weighted sum of the four components of  $H_{t+1}^{1/2} e_{t+1}$  and finally estimate the empirical quantile  $F_{w_s^\top \xi_{t+1}}^{-1}(\alpha)$ .

We estimate both VAR and each of the 11 MGARCH specifications on 495 windows of 500 observations each. Taking into account that each window begins five days later than the previous one, after having estimated the parameters on a particular window, we then use them to make a 5-day-ahead forecast. For each such forecast the parameters are taken constant, while the information on futures returns is updated every day. Because of the history needed for the first window estimation, the forecast time period is reduced to 2475 days.

For every model and every considered futures portfolio with weights  $w = (w_1, \dots, w_{28})^\top$  we estimate the exceedance rate  $\hat{\alpha}$ , which is the share of the observations for which the actual portfolio return is lower than the corresponding Value-at-Risk forecast:

$$\hat{\alpha}_w \stackrel{\text{def}}{=} n^{-1} \sum_{t=1}^n \mathbf{I} \left\{ \mathcal{R}_t < \widehat{VaR}_t(\alpha) \right\}, \quad (27)$$

where  $n = 2475$  is the number of forecast values for each portfolio. The relative deviation of  $\hat{\alpha}_w$  from the true  $\alpha$  is:

$$d_w \stackrel{\text{def}}{=} \frac{\hat{\alpha}_w - \alpha}{\alpha}.$$

In order to check the significance of the differences between  $\hat{\alpha}$  and  $\alpha$ , we employ the coverage test of Kupiec (1995) with the test statistic of the form:

$$K = 2 \log \left\{ \left( \frac{1 - \hat{\alpha}}{1 - \alpha} \right)^{n - \mathbf{I}(\alpha)} \left( \frac{\hat{\alpha}}{\alpha} \right)^{\mathbf{I}(\alpha)} \right\}, \quad (28)$$

where  $\mathbf{I}(\alpha) = \sum_{t=1}^n \mathbf{I} \mathcal{R}_t < \widehat{VaR}_t(\alpha)$ . The test statistic follows  $\chi^2(1)$  under  $H_0: \hat{\alpha} = \alpha$ .

In order to check the robustness of the procedure with respect to the choice of the portfolio, we run the study on a set  $W$  of  $|W| = 1000$  portfolios  $w^1, \dots, w^{1000}$ , where each portfolio  $w^q$  consists of  $p$  components with weights  $\{w_1^q, \dots, w_p^q\}$ ,  $p = 28$ ,  $q = 1, \dots, |W|$  being uniformly distributed over the cone  $S^p = \{(y_1, \dots, y_p)^\top \mid \sum_{i=1}^p y_i = 1\}$ . For weights generation we use Procedure 4 from Wang and Zionts (2006). By design, this procedure can produce only positive weights, but in our case it is in fact also desirable to allow for negative weights which would correspond to taking short positions in contracts. Indeed, unlike short positions in stocks, short positions in futures do not cause any additional transaction costs as compared to long positions. This is an argument in support of equal treatment of long and short positions both of which are likely to be taken, e.g. by a trader seeking arbitrage opportunities arising from futures prices deviating too far from some theoretical relationship. Moreover, higher variance of the weights will allow us to check if a particular model can predict not only the left tail of the shocks distribution correctly, but also its other intervals. Hence the weights were allowed to take negative values as well and the procedure had to be modified. Details are given in Appendix A.

We fix the first portfolio to be an equally-weighted portfolio,  $w_i^1 = 1/28 \approx 0.0357$ ,  $i = 1, \dots, 28$ . For the evaluation of the model performance on the whole portfolio set  $W$  we use average exceedance rates and average p-values of the Kupiec test measured across all  $|W|$  portfolios. Additionally we calculate  $A_W$ , which is the average relative deviation of  $d_w$ , and its standard error  $D_W$ :

$$A_W = \frac{1}{|W|} \sum_{w \in W} d_w, \quad D_W = \left\{ \frac{1}{|W|} \sum_{w \in W} (d_w - A_W)^2 \right\}^{1/2}. \quad (29)$$

Table 2 shows the results of the Value-at-Risk backtesting for the equally-weighted portfolio of 28 contracts and Table 3 summarises the results of the Value-at-Risk backtesting for all 1000 portfolios.

As can be seen from Table 2, for  $\alpha = 10\%$  and  $\alpha = 5\%$ ,  $H_0$  of the Kupiec test is not rejected at the 5% significance level for all models except univariate GARCH, which implies that treating shocks series as independent leads to severe underestimation of possible losses. However, the best performers for  $\alpha = 10\%$  and  $\alpha = 5\%$  are, respectively, BDECO with  $s_2$  and DECO while HAC-based models are second-best for these levels of  $\alpha$ . Only for  $\alpha = 1\%$  does a HAC-based model (Clayton HAC DCC with  $s_1$ ) outperform other benchmarks. Moreover, this model is one of the only two for which the hypothesis  $\hat{\alpha} = 1\%$  is not rejected at the 5% significance level. It is necessary to emphasise that while best performing models were able to forecast Value-at-Risk fairly well for  $\alpha = 10\%$  and  $\alpha = 5\%$ , the Value-at-Risk forecasts for  $\alpha = 1\%$  exhibit a relatively high exceedance rate, which means that the fat-tail distributions of the portfolio returns are not fully captured by the considered models. Overall results of the Value-at-Risk backtesting for the equally-weighted portfolio cannot be seen as satisfactory with regard to our expectations.

However, the situation is absolutely different on the aggregate level, i.e. for 1000 simulated portfolios on average. For all three  $\alpha$  values, the HAC-based DCC models outperform the benchmarks. They do not only produce the most accurate Value-at-Risk forecasts on average, as demonstrated by their exceedance rates, which are closest to the respective required  $\alpha$  values, but also generate these accurate forecasts on a regular basis, as can be seen from their low  $A_W$  and  $D_W$  values. It is also a very encouraging result that for two of the

Model	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 1\%$
DCC	9.616(0.522)	4.889(0.799)	1.495(0.021)
Gumbel AC DCC	9.778(0.712)	5.253(0.567)	1.455(0.033)
Clayton AC DCC	9.657(0.567)	4.848(0.728)	1.455(0.033)
Gumbel HAC DCC with $s_1$	9.697(0.614)	5.131(0.765)	1.495(0.021)
Clayton HAC DCC with $s_1$	9.657(0.567)	4.687(0.470)	<b>1.374(0.077)</b>
Gumbel HAC DCC with $s_2$	9.778(0.712)	5.131(0.765)	1.414(0.051)
Clayton HAC DCC with $s_2$	9.737(0.662)	4.889(0.799)	1.495(0.021)
DECO	9.778(0.712)	<b>5.051(0.908)</b>	1.495(0.021)
BDECO with $s_1$	9.697(0.614)	4.848(0.728)	1.455(0.033)
BDECO with $s_2$	<b>10.141(0.815)</b>	5.333(0.451)	1.616(0.005)
Univariate GARCH	16.283(0.000)	11.071(0.000)	4.162(0.000)

Table 2: Value-at-Risk backtesting results for the equally-weighted portfolio:  $\hat{\alpha}$  (in %) and Kupiec test p-values (in brackets). Results of the models yielding highest p-values are shown in bold.

three  $\alpha$  values, standard DCC performs worse than DECO in most cases. This means that a copula assumption can increase the forecasting qualities of a poorly-performing model significantly, so that it can overtake other benchmarks. The conclusion, valid for both equally-weighted portfolio and on the aggregate level, is that for  $\alpha = 1\%$  all models generally perform significantly worse, except for Clayton HAC DCC with  $s_1$ , which still shows a fair result (on the aggregate level it is one of the two models for which Kupiec test's null hypothesis is not rejected at the 10% significance level). The classical VaR exceedance plot is presented in Figure 13. Another perspective of the Value-at-Risk forecasts can be obtained by plotting their kernel densities over the whole forecasting period for each model and each  $\alpha$ . Figure 12 shows the left tail of the Value-at-Risk kernel density for  $\alpha = 1\%$  for each of the 11 considered models. The kernel densities are evaluated using the normal kernel smoother and the optimal bandwidth is estimated using Silverman's rule of thumb, as in section 4. It is easy to see that the models based on the Clayton copula are more inclined to provide low Value-at-Risk estimates than other models which is demonstrated

by the fatter left tails of the respective kernel densities. This is in line with our conclusion from the analysis of Table 2. BDECO with  $s_2$  performs worse and has higher exceedance rates than all other models except univariate GARCH (see Table 2), hence, as may be anticipated, it generates relatively thin left tails of predicted Value-at-Risk kernel density, see Figure 12. Summing up, one can conclude that as expected, HAC DCC models can be

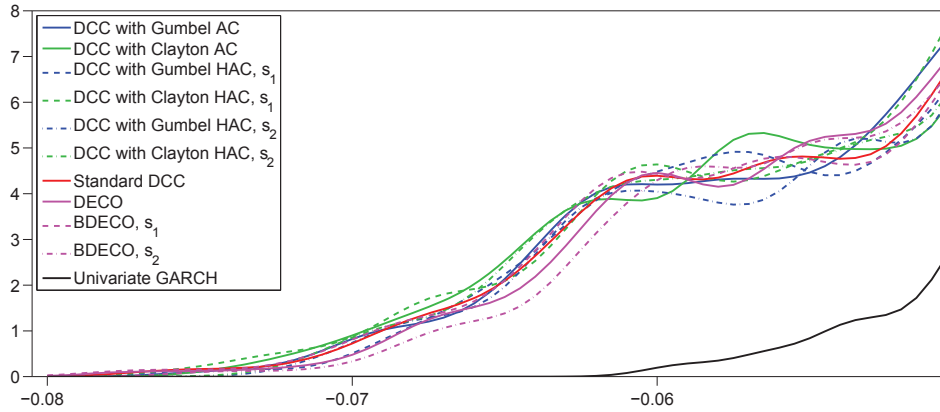


Figure 12: Kernel density of the value-at-risk ( $\alpha = 0.01$ ) forecast.

a very useful tool for risk management purposes. On average, they generate more accurate Value-at-Risk forecasts for various futures portfolios than nested models with stricter assumptions and even some other benchmark models (DECO).

## 6. Conclusions

This paper addresses an issue that can be of significant importance to many agents involved in commodity trading. The study models the dynamics of the heating oil and natural gas forward curves within one model. Multi-stage analysis of a large set of futures prices is carried out. For the analysis of the variance-covariance structure of the vector of the random shocks component, HAC-MGARCH models were used. This analysis allows us to forecast the distribution of the returns of any portfolio composed of the available futures contracts for short time periods. As shown in the study, Value-at-Risk estimates

Model	$\alpha = 10\%$			$\alpha = 5\%$			$\alpha = 1\%$		
	$\bar{\alpha}$ (p-value)	$A_W(D_W)$	$\bar{\alpha}$ (p-value)	$A_W(D_W)$	$\bar{\alpha}$ (p-value)	$A_W(D_W)$	$\bar{\alpha}$ (p-value)	$A_W(D_W)$	$\bar{\alpha}$ (p-value)
DCC	9.760(0.528)	-0.024(0.045)	5.051(0.476)	0.010(0.087)	1.450(0.068)	0.450(0.127)			
Gumbel AC DCC	9.806(0.557)	-0.019(0.044)	5.146(0.449)	0.029(0.089)	1.471(0.064)	0.471(0.139)			
Clayton AC DCC	9.678(0.495)	-0.032(0.045)	5.029(0.488)	0.006(0.086)	1.400(0.107)	0.400(0.132)			
Gumbel HAC DCC with $s_1$	9.853(0.571)	-0.015(0.044)	5.147(0.449)	0.029(0.088)	1.440(0.072)	0.440(0.122)			
Clayton HAC DCC with $s_1$	9.677(0.488)	-0.032(0.046)	4.981(0.475)	-0.004(0.089)	<b>1.362(0.141)</b>	<b>0.362(0.131)</b>			
Gumbel HAC DCC with $s_2$	<b>9.965(0.615)</b>	<b>-0.004(0.041)</b>	5.255(0.470)	0.051(0.073)	1.501(0.039)	0.501(0.116)			
Clayton HAC DCC with $s_2$	9.800(0.586)	-0.020(0.040)	<b>5.021(0.519)</b>	<b>0.004(0.077)</b>	1.449(0.062)	0.449(0.113)			
DECO	9.961(0.598)	-0.004(0.045)	5.251(0.493)	0.050(0.076)	1.560(0.027)	0.560(0.132)			
BDECO with $s_1$	9.869(0.572)	-0.013(0.046)	5.149(0.515)	0.030(0.079)	1.481(0.055)	0.481(0.136)			
BDECO with $s_2$	10.111(0.563)	0.011(0.047)	5.367(0.396)	0.073(0.079)	1.575(0.023)	0.575(0.131)			
Univariate GARCH	15.318(0.001)	0.532(0.123)	9.886(0.001)	0.977(0.219)	3.807(0.000)	2.807(0.711)			

Table 3: Summary of the Value-at-Risk backtesting results for 1000 portfolios:  $\bar{\alpha}$  is the average exceedance rates across all portfolios (in %), p-values are average Kupiec test p-value across all portfolios.  $A_W$  and  $D_W$  are defined as in (29). Results of the models yielding highest average p-values are shown in bold.



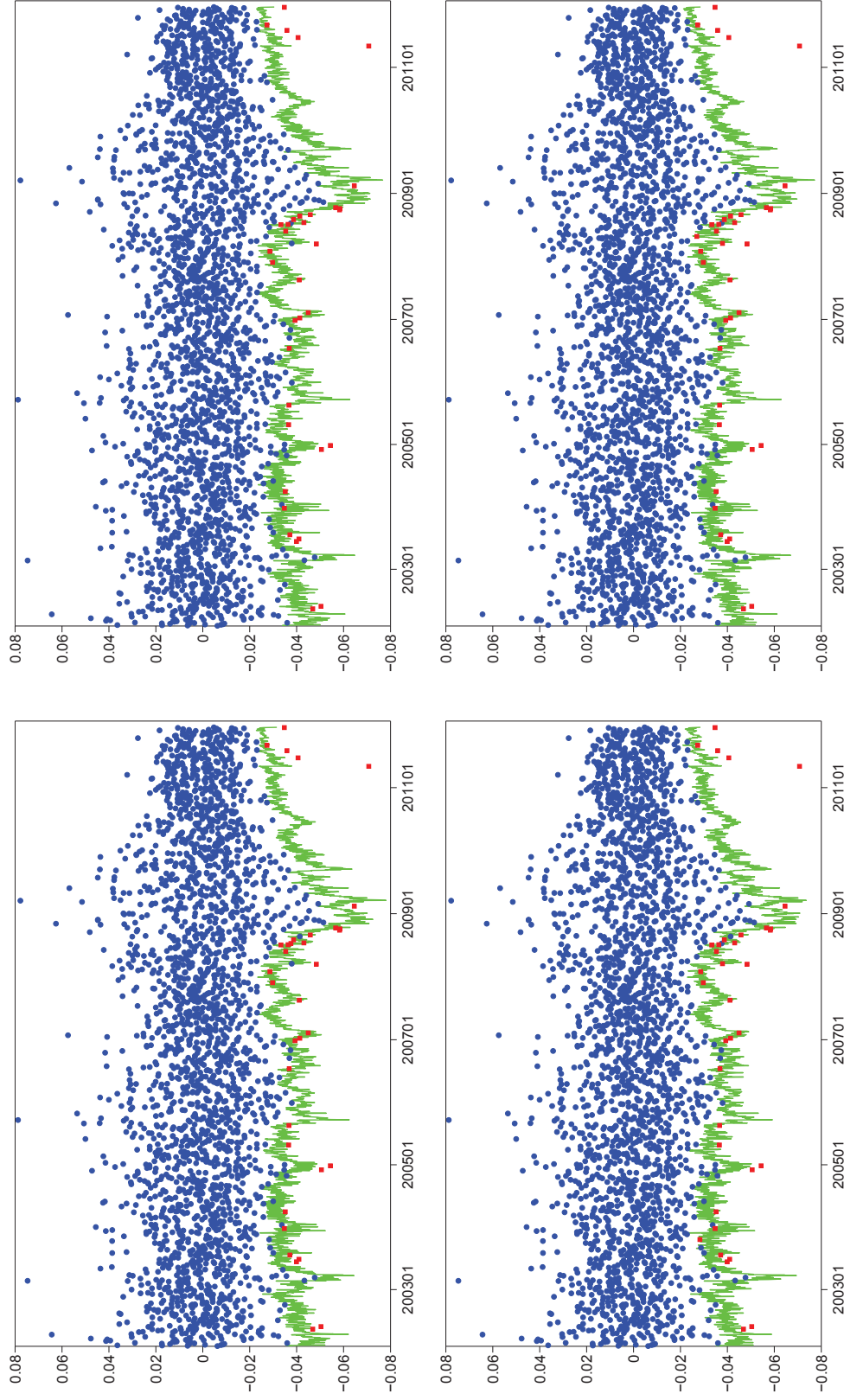


Figure 13: Value-at-Risk ( $\alpha = 1\%$ ) exceedance plots for standard DCC (upper left), DCC with Gumbel AC (upper right), DCC with Gumbel HAC,  $s_2$  (lower left) and BDECO,  $s_2$  (lower right) (6.02.2002 – 19.12.2011). Data series: predicted Value-at-Risk (green line), portfolio returns higher than predicted Value-at-Risk (blue dots) and portfolio returns lower than predicted Value-at-Risk (red squares).

derived from the forecasts produced by HAC DCC models are accurate, and these models outperform other benchmark models on a consistent basis, as shown by the Value-at-Risk backtesting procedure carried out on a set of futures portfolios.

The research can be extended along several directions. First, one can further combine copulae with different MGARCH models, e.g. employ copula-based DECO models. Second, other copula types, such as vine copulae, can be considered. Furthermore, more sophisticated methods to the determination of the estimation time window, such as local adaptive methods, can be applied, which can lead to a better understanding of the time evolution of the processes and more precise forecasts.

## Appendix A. Algorithm used to generate random weights

1. Generate a set  $\Lambda_1$  of  $|\Lambda_1| = 1000$  portfolios of  $p = 28$  components with the weights of each component in each portfolio  $\lambda_1^{1,q}, \dots, \lambda_p^{1,q}$ ,  $q = 1, \dots, |\Lambda_1|$  uniformly distributed over the cone  $S^p = \{(y_1, \dots, y_p)^\top | y_i \geq 0, \sum_{i=1}^p y_i = 1\}$ . Multiply all weights in all portfolios by 2.
2. Generate a set  $\Lambda_2$  of  $|\Lambda_2| = 1000$  portfolios of  $p = 28$  components with the weights of each component in each portfolio  $\lambda_1^{2,q}, \dots, \lambda_p^{2,q}$ ,  $q = 1, \dots, |\Lambda_1|$  uniformly distributed over the cone  $S^p = \{(y_1, \dots, y_p)^\top | y_i \geq 0, \sum_{i=1}^p y_i = 1\}$ .
3. Let  $W$  be a set of  $|W| = 1000$  portfolios with weights calculated as  $w_i^q = \lambda_i^{1,q} - \lambda_i^{2,q}$  for all  $i = 1, \dots, p$  and  $q = 1, \dots, |W|$ .
4. The mean weight  $\bar{w}$  in the resulting portfolios set  $W$  is equal to  $\bar{w} = 1/28 = 0.0357$  and the weight's standard deviation  $\sigma_w = 0.077$  (measured across all contracts and portfolios).

Annual Energy Outlook, June 2012. U.S. Energy Information Administration. Report DOE/EIA-0383(2012), [www.eia.gov/forecasts/aeo/](http://www.eia.gov/forecasts/aeo/).

Asche, F., Osmundsen, P., Sandsmark, M., 2006. The UK market for natural gas, oil and electricity: are prices decoupled? *The Energy Journal* 27 (2), 27–40.

- Bachmeier, L., Griffin, J., 2006. Testing for market integration: crude oil, coal, and natural gas. *The Energy Journal* 27 (2), 55–72.
- Chambers, M. J., Bailey, R. E., 1996. A theory of commodity price fluctuations. *Journal of Political Economics* 104 (5), 924–957.
- Chantziara, T., Skiadopoulos, G., 2008. Can the dynamics of the term structure of petroleum futures be forecasted? Evidence from major markets. *Energy Economics* 30, 962–985.
- Cortazar, G., Schwartz, E. S., Summer 1994. The valuation of commodity-contingent claims. *The Journal of Derivatives*, 27–39.
- Cortazar, G., Schwartz, E. S., 2003. Implementing a stochastic model for oil futures prices. *Energy Economics* 25, 215–238.
- Dawson, P. J., Sanjuán, A. I., White, B., 2006. Structural breaks and the relationship between barley and wheat futures prices on the London International Financial Futures Exchange. *Review of Agricultural Economics* 28 (4), 585–594.
- De Bock, R., Gijón, J., June 2011. Will natural gas prices decouple from oil prices across the pond? Working Paper WP/11/143, International Monetary Fund.
- Engle, R., 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* 20 (3), 339–350.
- Engle, R., Granger, C., 1987. Co-integration and error correction: Representation, estimation and testing. *Econometrica* 55 (2), 251–276.
- Engle, R., Kelly, R., 2012. Dynamic equicorrelation. *Journal of Business & Economic Statistics* 30 (2), 212–228.
- Engle, R., Kroner, K., 1995. Multivariate simultaneous generalized ARCH. *Econometric Theory* 11, 122–150.

- Eydeland, A., Geman, H., 1998. Pricing power derivatives. *RISK* September, 71–73.
- Gabillon, J., 1991. The term structures of oil futures prices. Working Paper WPM 17, Oxford Institute for Energy Studies.
- Grasso, M., Manera, M., 2007. Asymmetric error correction models for the oil-gasoline price relationship. *Energy Policy* 35, 156–177.
- Hartley, P., Medlock, K., Rosthal, J., 2008. The relationship of natural gas to oil prices. *The Energy Journal* 29 (3), 47–66.
- Heath, D., Jarrow, R. A., Morton, A., 1990. Contingent claim valuation with a random evolution of interest rates. *The Review of Futures Markets* 9 (1), 54–76.
- Hoeffding, W., 1940. Maßtabinvariante Korrelationstheorie, *Schriften des Mathematischen Instituts und des Instituts für angewandte Mathematik der Universität Berlin*, 5, Heft 3, 179–233. [Reprinted as Scale-invariant correlation theory in *The Collected Works of Wassily Hoeffding*, N.I. Fisher and P.K. Sen editors, Springer-Verlag, New York, 57–107].
- Hofert, M., 2008. Sampling Archimedean copulas. *Computational Statistics and Data Analysis* 52, 5163–5174.
- Jin, H. J., 2007. Heavy-tailed behaviour of commodity price distribution and optimal hedging demand. *The Journal of Risk and Insurance* 74 (4), 863–881.
- Koekebakker, S., Ollmar, F., 2005. Forward curve dynamics in the Nordic electricity market. *Managerial Finance* 31 (6), 73–94.
- Kullback, L., Leibler, R., 1951. On information and sufficiency. *Annals of Mathematical Statistics* 22, 79–86.
- Kupiec, P., 1995. Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives* 3, 73–84.

- Lee, T.-H., Long, X., 2009. Copula-based multivariate GARCH model with uncorrelated dependent errors. *Journal of Econometrics* 150, 207–218.
- Litterman, R., Scheinkman, J., June 1991. Common factors affecting bond returns. *The Journal of Fixed Income*, 54–61.
- Liu, P., Tang, K., 2010. No-arbitrage conditions for storable commodities and the modeling of futures term structures. *Journal of Banking & Finance* 34, 1675–1687.
- Lütkepohl, H., Krätzig, M. (Eds.), 2004. *Applied Time Series Econometrics. Themes in Modern Econometrics*. Cambridge University Press.
- McNeil, A., 2008. Sampling nested Archimedean copulas. *Journal of Statistical Computation and Simulation* 78 (6), 567–581.
- McNeil, A., Nešlehová, J., 2009. Multivariate Archimedean copulas,  $d$ -monotone functions and  $l_1$ -norm symmetric distributions. *The Annals of Statistics* 37 (5B), 3059–3097.
- Ohana, S., 2010. Modeling global and local dependence in a pair of commodity forward curves with an application to the US natural gas and heating oil markets. *Energy Economics* 32, 373–388.
- Okhrin, O., Okhrin, Y., Schmid, W., 2012a. On the structure and estimation of hierarchical Archimedean copulas. *Journal of Econometrics*, under revision.
- Okhrin, O., Okhrin, Y., Schmid, W., 2012b. Properties of hierarchical Archimedean copulas. *Statistics and Risk Modelling*, forthcoming.
- Onour, I., April 2009. Natural gas markets: How sensitive to crude oil price changes? MPRA Paper 14937, Arab Planning Institute, Munich Personal RePEc Archive.  
URL <http://mpra.ub.uni-muenchen.de/14937/>
- Pilipovic, D., 2007. *Energy risk: Valuing and managing energy derivatives*, 2nd Edition. McGraw-Hill.

- Reisman, H., 1991. Movements of the term structure of commodity futures and the pricing of commodity claims.
- Schwartz, E. S., 1997. The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance* 52 (3), 923–973.
- Sklar, A., 1959. Fonctions de répartition à  $n$  dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris* 8, 229–231.
- Tolmasky, C., Hindanov, D., 2002. Principal components analysis for correlated curves and seasonal commodities: The case of the petroleum market. *The Journal of Futures Markets* 22 (11), 1019–1035.
- Tse, Y., Tsui, A., 2002. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business and Economic Statistics* 20, 351–362.
- Wang, J., Zionts, S., 2006. Random weight generation in multiple criteria decision models. *MCDM 2006, Chania, Greece, June 19-23, 2006*.

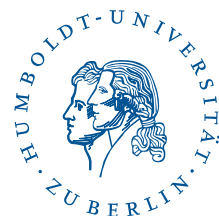
# SFB 649 Discussion Paper Series 2012

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "HMM in dynamic HAC models" by Wolfgang Karl Härdle, Ostap Okhrin and Weining Wang, January 2012.
- 002 "Dynamic Activity Analysis Model Based Win-Win Development Forecasting Under the Environmental Regulation in China" by Shiyi Chen and Wolfgang Karl Härdle, January 2012.
- 003 "A Donsker Theorem for Lévy Measures" by Richard Nickl and Markus Reiß, January 2012.
- 004 "Computational Statistics (Journal)" by Wolfgang Karl Härdle, Yuichi Mori and Jürgen Symanzik, January 2012.
- 005 "Implementing quotas in university admissions: An experimental analysis" by Sebastian Braun, Nadja Dwenger, Dorothea Kübler and Alexander Westkamp, January 2012.
- 006 "Quantile Regression in Risk Calibration" by Shih-Kang Chao, Wolfgang Karl Härdle and Weining Wang, January 2012.
- 007 "Total Work and Gender: Facts and Possible Explanations" by Michael Burda, Daniel S. Hamermesh and Philippe Weil, February 2012.
- 008 "Does Basel II Pillar 3 Risk Exposure Data help to Identify Risky Banks?" by Ralf Sabiwalsky, February 2012.
- 009 "Comparability Effects of Mandatory IFRS Adoption" by Stefano Cascino and Joachim Gassen, February 2012.
- 010 "Fair Value Reclassifications of Financial Assets during the Financial Crisis" by Jannis Bischof, Ulf Brüggemann and Holger Daske, February 2012.
- 011 "Intended and unintended consequences of mandatory IFRS adoption: A review of extant evidence and suggestions for future research" by Ulf Brüggemann, Jörg-Markus Hitz and Thorsten Sellhorn, February 2012.
- 012 "Confidence sets in nonparametric calibration of exponential Lévy models" by Jakob Söhl, February 2012.
- 013 "The Polarization of Employment in German Local Labor Markets" by Charlotte Senftleben and Hanna Wielandt, February 2012.
- 014 "On the Dark Side of the Market: Identifying and Analyzing Hidden Order Placements" by Nikolaus Hautsch and Ruihong Huang, February 2012.
- 015 "Existence and Uniqueness of Perturbation Solutions to DSGE Models" by Hong Lan and Alexander Meyer-Gohde, February 2012.
- 016 "Nonparametric adaptive estimation of linear functionals for low frequency observed Lévy processes" by Johanna Kappus, February 2012.
- 017 "Option calibration of exponential Lévy models: Implementation and empirical results" by Jakob Söhl und Mathias Trabs, February 2012.
- 018 "Managerial Overconfidence and Corporate Risk Management" by Tim R. Adam, Chitru S. Fernando and Evgenia Golubeva, February 2012.
- 019 "Why Do Firms Engage in Selective Hedging?" by Tim R. Adam, Chitru S. Fernando and Jesus M. Salas, February 2012.
- 020 "A Slab in the Face: Building Quality and Neighborhood Effects" by Rainer Schulz and Martin Wersing, February 2012.
- 021 "A Strategy Perspective on the Performance Relevance of the CFO" by Andreas Venus and Andreas Engelen, February 2012.
- 022 "Assessing the Anchoring of Inflation Expectations" by Till Strohsal and Lars Winkelmann, February 2012.

**SFB 649, Spandauer Straße 1, D-10178 Berlin**  
**<http://sfb649.wiwi.hu-berlin.de>**

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".





# SFB 649 Discussion Paper Series 2012

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 023 "Hidden Liquidity: Determinants and Impact" by Gökhan Cebiroglu and Ulrich Horst, March 2012.
- 024 "Bye Bye, G.I. - The Impact of the U.S. Military Drawdown on Local German Labor Markets" by Jan Peter aus dem Moore and Alexandra Spitz-Oener, March 2012.
- 025 "Is socially responsible investing just screening? Evidence from mutual funds" by Markus Hirschberger, Ralph E. Steuer, Sebastian Utz and Maximilian Wimmer, March 2012.
- 026 "Explaining regional unemployment differences in Germany: a spatial panel data analysis" by Franziska Lottmann, March 2012.
- 027 "Forecast based Pricing of Weather Derivatives" by Wolfgang Karl Härdle, Brenda López-Cabrera and Matthias Ritter, March 2012.
- 028 "Does umbrella branding really work? Investigating cross-category brand loyalty" by Nadja Silberhorn and Lutz Hildebrandt, April 2012.
- 029 "Statistical Modelling of Temperature Risk" by Zografia Anastasiadou, and Brenda López-Cabrera, April 2012.
- 030 "Support Vector Machines with Evolutionary Feature Selection for Default Prediction" by Wolfgang Karl Härdle, Dedy Dwi Prastyo and Christian Hafner, April 2012.
- 031 "Local Adaptive Multiplicative Error Models for High-Frequency Forecasts" by Wolfgang Karl Härdle, Nikolaus Hautsch and Andrija Mihoci, April 2012.
- 032 "Copula Dynamics in CDOs." by Barbara Choroś-Tomczyk, Wolfgang Karl Härdle and Ludger Overbeck, May 2012.
- 033 "Simultaneous Statistical Inference in Dynamic Factor Models" by Thorsten Dickhaus, May 2012.
- 034 "Realized Copula" by Matthias R. Fengler and Ostap Okhrin, Mai 2012.
- 035 "Correlated Trades and Herd Behavior in the Stock Market" by Simon Jurkatis, Stephanie Kremer and Dieter Nautz, May 2012
- 036 "Hierarchical Archimedean Copulae: The HAC Package" by Ostap Okhrin and Alexander Ristig, May 2012.
- 037 "Do Japanese Stock Prices Reflect Macro Fundamentals?" by Wenjuan Chen and Anton Velinov, May 2012.
- 038 "The Aging Investor: Insights from Neuroeconomics" by Peter N. C. Mohr and Hauke R. Heekeren, May 2012.
- 039 "Volatility of price indices for heterogeneous goods" by Fabian Y.R.P. Bocart and Christian M. Hafner, May 2012.
- 040 "Location, location, location: Extracting location value from house prices" by Jens Kolbe, Rainer Schulz, Martin Wersing and Axel Werwatz, May 2012.
- 041 "Multiple point hypothesis test problems and effective numbers of tests" by Thorsten Dickhaus and Jens Stange, June 2012
- 042 "Generated Covariates in Nonparametric Estimation: A Short Review." by Enno Mammen, Christoph Rothe, and Melanie Schienle, June 2012.
- 043 "The Signal of Volatility" by Till Strohsal and Enzo Weber, June 2012.
- 044 "Copula-Based Dynamic Conditional Correlation Multiplicative Error Processes" by Taras Bodnar and Nikolaus Hautsch, July 2012

**SFB 649, Spandauer Straße 1, D-10178 Berlin**  
**<http://sfb649.wiwi.hu-berlin.de>**

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".





## SFB 649 Discussion Paper Series 2012

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 045 "Additive Models: Extensions and Related Models." by Enno Mammen, Byeong U. Park and Melanie Schienle, July 2012.
- 046 "A uniform central limit theorem and efficiency for deconvolution estimators" by Jakob Söhl and Mathias Trabs, July 2012
- 047 "Nonparametric Kernel Density Estimation Near the Boundary" by Peter Malec and Melanie Schienle, August 2012
- 048 "Yield Curve Modeling and Forecasting using Semiparametric Factor Dynamics" by Wolfgang Karl Härdle and Piotr Majer, August 2012
- 049 "Simultaneous test procedures in terms of p-value copulae" by Thorsten Dickhaus and Jakob Gierl, August 2012
- 050 "Do Natural Resource Sectors Rely Less on External Finance than Manufacturing Sectors? " by Christian Hattendorff, August 2012
- 051 "Using transfer entropy to measure information flows between financial markets" by Thomas Dimpfl and Franziska J. Peter, August 2012
- 052 "Rethinking stock market integration: Globalization, valuation and convergence" by Pui Sun Tam and Pui I Tam, August 2012
- 053 "Financial Network Systemic Risk Contributions" by Nikolaus Hautsch, Julia Schaumburg and Melanie Schienle, August 2012
- 054 "Modeling Time-Varying Dependencies between Positive-Valued High-Frequency Time Series" by Nikolaus Hautsch, Ostap Okhrin and Alexander Ristig, September 2012
- 055 "Consumer Standards as a Strategic Device to Mitigate Ratchet Effects in Dynamic Regulation" by Raffaele Fiocco and Roland Strausz, September 2012
- 056 "Strategic Delegation Improves Cartel Stability" by Martijn A. Han, October 2012
- 057 "Short-Term Managerial Contracts and Cartels" by Martijn A. Han, October 2012
- 058 "Private and Public Control of Management" by Charles Angelucci and Martijn A. Han, October 2012
- 059 "Cartelization Through Buyer Groups" by Chris Doyle and Martijn A. Han, October 2012
- 060 "Modelling general dependence between commodity forward curves" by Mikhail Zolotko and Ostap Okhrin, October 2012

**SFB 649, Spandauer Straße 1, D-10178 Berlin**  
**<http://sfb649.wiwi.hu-berlin.de>**

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

